Notes:
1. Untranslatable words are replaced with asterisks (****).
2. Texts in the figures are not translated and shown as it is.

Translated: 04:49:03 JST 06/09/2006
Dictionary: Last updated 05/26/2006 / Priority: 1. Information communication technology (ICT) / 2. Electronic engineering / 3. JIS (Japan Industrial Standards) term

---

FULL CONTENTS

[Claim(s)]

[Claim 1] Two or more decision-making subjects exist, and a decision-making subject has two or more action alternative, respectively. In the task request in two or more decision-making subject environment and the task undertaking action determination method of determining each one of gains by combination of action which each decision-making subject took, and determining action that the gain will become large [without describing the combination of action, and the relation of gain comprehensively to all the states of a decision-making subject, when describing the gain to the combination of action] Express the method of change of the gain by action using a change rule, and problem expression is made simple by calculating and drawing the combination of the action to a future state, and the relation of gain using the change rule if needed. Moreover, which action should be taken and when calculating, carry out backward reasoning from future action, and action is not determined. By evaluating the change condition of the future gain which the action in this time causes, and including the evaluation in the decision criterion of selection of operation, without determining action only from short-term profit and loss at present The task request and the task undertaking action determination method in two or more decision-making subject environment which are characterized by making the long-term gain of each decision-making subject increase.

[Claim 2] In the task request in two or more decision-making subject environment and the task undertaking action determination method in the system to which two or more computers were connected through the network The process which calculates the gain to action after this time from the change rule of gain and gain over the action in this time of the decision-making subject stored in the storage of a computer, The task request and the task undertaking action determination method in two or more decision-making subject environment which are characterized by having the process which determines action of a decision-making subject that the use calculated from the gain by action of the decision-making subject in this time and the

future gain change caused by the action will serve as the maximum.

[Detailed Description of the Invention]

[0001]

[Field of the Invention] As opposed to the decision problem of the task request on the network system of the computer by which interconnection of this invention was carried out, and task undertaking When action of whether a computer with the task which should be processed performs a data transfer request and the reference request of a database to other computers, or to process a task by self is determined or there is a task request from other computers It is related with the task request in two or more decision-making subject environment and the task undertaking action determination method of determining whether it is taken over and that action which is not taken over.

[0002]

[Description of the Prior Art] Two or more computers are connected to a network, each computer has a task, and the environment where information is exchanged between computers for tasking is considered. Suppose that there is also no main computer direct that one task takes cooperative action mutually rather than that is given [and] to the computer of these plurality on the whole. Here, suppose that the data in the database which Computer B manages is effective in tasking of Computer A, for example. However, even when he has none of this data, tasking presupposes that it is possible. At this time, there is selection of whether Computer A requests data reference from Computer B, and performs tasking or for it not to request but to perform tasking by self. Moreover, Computer B has the answering selection which responds to it, when there is a request.

[0003] The action which generally serves as action used as self profits and a partner's profits is not always in agreement. That is, it is impossible to force a partner computer into action advantageous to self. Therefore, a partner computer must choose the action which makes self profits the maximum under assumption of acting rationally. Thus, there is a game theory as a method of dealing with an action selection problem in case there is a partner.

[0004] When the decision-making subject of (1) plurality exists, a game theory has two or more alternative of respectively action of (2) decision-making subjects and (3) each decision-making subject chooses action, With the combination of the action, each gain is determined and a decision-making subject besides (4) gives the method of being the basis of the premise of acting so that all the members and a use (gain) may be maximized, and determining which action should be chosen. Moreover, the method of analyzing how change of a setup of a use changing the property of the whole system is given (refer to [Mitsuo Suzuki and "new game

theory" Keiso-Shobo, 1994], and [the Toshio Nishida and "theory-of-game" Union of Japanes Scientists and Engineers and 1973]).

[0005] The process of acquisition of action selection and the gain by it is called a game, and a sense resolution subject is called a participant. Moreover, each stage which constitutes a game, i.e., each action selection opportunity, is called a move. Here, the number of participants deals with the game of noncooperation and complete information by 2. It says that the agreement of the beforehand which has legal force as it is noncooperative does not exist. It is completely unrelated whether it seems to have taken the concerted action as a result or it is not visible. Moreover; complete information means that all the participants know all actions chosen in the process before it.

[0006] There is expression by a game tree as one of the methods showing a game. The example of expression of the game by a game tree is shown in drawing 9 R> 9. PA and PB The move of the participants A and B in a game is expressed, respectively. A branch expresses selectable action. This figure expresses process in which B chooses action ofb1 orb2, after A chooses action of a1 or a2 first and looks at it. Each gain is decided by combination of Participants' A and B action, and it is shown at the endpoint. This is called a gain vector.

[0007] Now, a game has the game which a many steps [which both sides choose action 1 time respectively and not only finish, but repeat as a portion the game shown in drawing 9] hand follows. This is called multistage game. On these Descriptions, the period until both sides choose one hand of actions from this time at a time and gain is decided first is called a game at present as one unit.

[0008] Now, when the game tree from which the total of a move is set to N is given, the conventional method of making action selection by each move is shown below. Here, it is defined as utility = gain and let N be even number.

[0009] (The action selection method on a game tree)

(1) Consider it as i<-N.

(2) the action which looks at a utility vector and makes the use of the participant in this move the maximum about each branch point Of Move i -- each branch point -- carry out selection behavior. Here, the utility vector is the same as that of a gain vector.

(3) About each branch point of a move i-1, choose the action which makes the use of the participant in this move the maximum from the utility vectors to the selection behavior of (2), and consider it as the selection behavior in each branch point.

(4) If it is i= 2, it will go to (7) as i<- (i-1).

(5) If it is not i= 2, the utility vector value over the selection behavior in the move i-1 which follows it will be applied to the utility vector to each action at the branch point i-2.

(6) Go to (2) as i<- (i-2).

(7) Let selection behavior in the branch point chosen by Move i be the selection behavior in this game. In a move 1, the number of the branch points is one and let them be the branch points which had it chosen.
(8) If it is i=N, it will end.
Let the branch point which follows the selection behavior of (9) and (7) be the selection branch point in the move i<- (i+1).
It goes to (10) and (7).
[0010] In a perfect information game, it can be known in which branch point each participant is, and the participant can draw the justification in the noncooperation game of the upper method at each branch point from the fact of choosing the action which always maximizes a self use henceforth [it].
[0011] If the example of drawing 9 is referred to, since the use of action b2 is larger than b1 for B, by a move 2 (PB), action b2 will be chosen at both of the branch points. Next, in a move 1 (PA), it becomes comparison of a1 (utility vector (0, 8)) and a2 (utility vector (2, 2)). For A, since the use is [a2] larger, it is chosen here. In this game, action sequence a2 ->b2 appear after all.
[0012]
[Problem to be solved by the invention] Now, man performs input process of a gain vector and a computer carries out to calculating action selection. The block diagram in the case of realizing using the conventional technology is shown in drawing 10 R> 0. When it is going to express a game as a game tree, the following problems arise.
[0013] (a) Generally the action taken at present changes the gain vector of the game performed after it. Therefore, it is required to be able to deal with the change. Treatment of carrying out by preparing only one game with a small size and repeating it is inadequate.
[0014] (b) Since a game tree will become huge if severalN of a move becomes large, it becomes difficult to write down on the form developed beforehand.
[0015] (c) If the group of a participant's condition and the gain vector of the game performed there is described, a game tree can be developed if needed. However, if a state number increases, the amount of description of the group of a state and a gain vector will become huge, and management will become difficult.
[0016] Moreover, when calculating, the following problems arise.
[0017] (d) Calculation of the conventional method about the above-mentioned action selection advances backward first from the last move N. However, since not every participant has the legal force which continues making a participant participate in a game, it cannot know the move of which will be performed beforehand. Therefore, when N is not given, it cannot be determined where calculation should be begun from.
[0018] (e) If N becomes large even when the total N of the move is known, a game tree will

become huge, computational complexity will increase, and it will come to require time by action determination.

[0019] (f) When giving up the calculation from the last move N and making an action decision about a game at present, future profit and loss cannot be included in calculation. Therefore, even if it loses in the short term, selection of the action which obtains gain in the long run becomes impossible.

[0020] [the place which this invention was made in view of the above, and is made into the purpose] While making decisions by introducing the description under the change rule of gain for a short time, selection of the action gained in the long run is attained, and it is in offering the task request in two or more decision-making subject environment and the task undertaking action determination method which may increase the gain which can be gained in the long run.

[0021]

[Means for solving problem] In order to attain the above-mentioned purpose, [this invention according to claim 1] Two or more decision-making subjects exist, and a decision-making subject has two or more action alternative, respectively. In the task request in two or more decision-making subject environment and the task undertaking action determination method of determining each one of gains by combination of action which each decision-making subject took, and determining action that the gain will become large [without describing the combination of action, and the relation of gain comprehensively to all the states of a decision-making subject, when describing the gain to the combination of action] Express the method of change of the gain by action using a change rule, and problem expression is made simple by calculating and drawing the combination of the action to a future state, and the relation of gain using the change rule if needed. Moreover, which action should be taken and when calculating, carry out backward reasoning from future action, and action is not determined. Let it be a summary to make the long-term gain of each decision-making subject increase by evaluating the change condition of the future gain which the action in this time causes, and including the evaluation in the decision criterion of selection of operation, without determining action only from short-term profit and loss at present.

[0022] If it is in this invention according to claim 1, the method of change of the gain by action is expressed using a change rule. [or] when calculating which action problem expression is made simple and should be taken by calculating and drawing the combination of the action to a future state, and the relation of gain using the change rule if needed The long-term gain of each decision-making subject is made to increase by evaluating the change condition of the future gain which the action in this time causes, and including the evaluation in the decision criterion of selection of operation.

[0023] Moreover, this invention according to claim 2 is set to the task request in two or more

decision-making subject environment and the task undertaking action determination method in the system to which two or more computers were connected through the network. The process which calculates the gain to action after this time from the change rule of gain and gain over the action in this time of the decision-making subject stored in the storage of a computer, Let it be a summary to have the process which determines action of a decision-making subject that the use calculated from the gain by action of the decision-making subject in this time and the future gain change caused by the action will serve as the maximum.

[0024] If it is in this invention according to claim 2, the gain to action after this time is calculated from the change rule of gain and gain over the action in this time of a decision-making subject. Action of a decision-making subject is determined that the use calculated from the gain by action of the decision-making subject in this time and the future gain change caused by the action will serve as the maximum.

[0025]

[Mode for carrying out the invention] The form of operation of this invention is hereafter explained using Drawings.

[0026] Drawing 1 is the block diagram showing the composition of the equipment which enforces the task request in two or more decision-making subject environment and the task undertaking action determination method concerning one embodiment of this invention. The equipment shown in this figure has the storage 1, I/O device 3, the arithmetic unit 5 that performs utility calculus, and the communication device 7 which communicates with a partner computer, and it differs in that added the change rule of the gain vector to the database of the game tree, and it was newly added to the storage 1 to the conventional equipment shown in drawing 10 . Thus, by newly adding the method of describing the change rule of a gain vector independently, expression by the game tree of a game can be made easy. If this performs a game at present and a certain action is performed, it will describe how a gain vector changes in the game after it.

[0027] Drawing 2 shows the example of description of the change rule over drawing 9 . Drawing 2 (a) is a change rule over A, and drawing 2 (b) is a change rule over B. Moreover, in the table of this figure, x shows the value of the gain in the game before it. When action called a1 and b1 is taken, the line a1 of a table and the intersection of a sequence b1 are seen. About A, it is with an equation x+1, x+2, x-1, and x+1, and corresponds to the formula of the gain of (a1, b1) in the next time, (a1, b2), (a2, b1), and (a2, b2), respectively. If it actually calculates, it can be found as follows.

[0028] (a1, b1) : x+1=5+1=6 (a1, b2) : x+2=0+2=2 (a2, b1) : x-1=8-1=7 (a2, b2) : x+1=2+1=3

[0029] Drawing 3 will be obtained if it calculates to other action groups or B similarly. The gain vector of subsequent games is also calculable if needed.

[0030] Next, the action selection method which added evaluation of the action in this time and

the relation of subsequent games is proposed. The action in this time is after it, and for self, this evaluates whether it leads to a more advantageous game, i.e., the game which may obtain bigger gain, or it does not lead, unites it and evaluation of the profit and loss on the game in this time, and performs action selection. The following formulas define a use.
[0031]
[Mathematical formula 1]
Utility = (gain of the game in this time)
+ ((w) Coefficient of expectation) x (gain change in subsequent games) (1)
The participant in a game chooses the action which makes this use the maximum. Coefficients of expectation differ for every participant, and are adjusted according to each task which it has.

[0032] The computational procedure of the proposal method is shown below. Here, it is careful of a game at present referring to a game until both sides choose one hand of actions at a time and gain is decided.
[0033] (The action selection method which incorporated evaluation of the game change after being based on action at present)
(1) Set a move at present to i.
(2) Calculate the use of an upper type (1) from the gain of a game at present.
(3) Let action which makes the use of the participant in this move the maximum be the selection behavior in each branch point about each branch point of a move i+1.
(4) About each branch point of Move i, choose the action which makes the use of the participant in this move the maximum from the utility vectors to the selection behavior of (3), and consider it as the selection behavior in this game.
(5) Let selection behavior in the branch point which follows the selection behavior in the branch point i be the selection behavior in a move i+1.
(6) End.
[0034] Reduction of the amount of description is expectable from description of only the group of a state and a gain vector with introduction of the expression under the above-mentioned change rule. If a game tree will be developed if needed, required storage capacity is reducible.
[0035] Moreover, the following things become possible by the action selection method which incorporated evaluation of the game change after being based on the above-mentioned action.

[0036] (a) Since the proposal technique does not perform strict calculation like the conventional technology but presumes the increment of future gain, it cannot necessarily perform the always right judgment. However, it becomes unnecessary to calculate the whole game tree comprehensively like the conventional technology, and computation time is reduced. This is effective especially when temporal restrictions, such as control of a physical

system, occur.

[0037] (b) If there is action which may obtain big profits in the future even if it stops being judgment only for profit and loss at present and loses at present even when the total N of a move is not known beforehand, a possibility of choosing it will arise.

[0038] (c) By operation of a coefficient of expectation, the suitable action selection according to the load in a participant's this time becomes possible. For example, if generous at present, the degree of expectation can be enlarged and it can respond to the high request of the load from a partner cooperatively. If a partner also uses the same action selection method, a self load is high, and a possibility that a partner will consent will increase to distribute a task to a partner. Moreover, if hard-pressed, a coefficient of expectation can be made small, the request from a partner can be refused, and it can concentrate on self tasking.

[0039] Next, a distribution-of-information problem is considered between Computer A and Computer B as a concrete example. Problem setting here is shown below.

(1) Computer A and Computer B are connected to the same network.

(2) Computer A advances the data reference demand of the database which B manages to the partner computer B for tasking.

(3) If a reference request is taken out to a partner, cost (negative gain) will generate Computer A.

(4) If Computer B accepts the database reference request of a partner computer, cost (negative gain) will generate it. If the demand of a partner computer is refused, cost will not start.

(5) Since it will be useful for tasking if data can be gained from a partner, Computer A can obtain positive gain.

(6) Computer A passes a packet unnecessary to a network at the time of tasking, and brings a load to a partner computer. The grade of a load is based on the knowledge level of Computer A. When the knowledge level about the distribution of information on a network is high, the load given to the circumference can be made small.

(7) If Computer B accepts a reference request, the knowledge about the distribution of information which accompanies it will be accumulated in Computer A, and the knowledge level of A will go up.

(8) When there is no exchange of information between computers, Computer A cannot acquire the knowledge about the distribution of information, but a knowledge level falls.

[0040] Action of this problem and the relation of gain are shown in drawing 4 . The cost about a question and a reply is shown in drawing 4 (a). The gain by data capture is shown in drawing 4 (b). The initial value of the cost produced in B along with tasking is shown in drawing 4 (c). The gain by performing this game once serves as the sum of (a), (b), and (c). 51 of drawing 5 R> 5 shows the game tree in an initial state. The cost of drawing 4 (c) changes with the knowledge

levels of Computer A. The change rule of the cost is shown in drawing 6 . 53 of drawing 5 shows the game tree in the next state of the initial state searched for using this rule.
[0041] It is problems here how Computers A and B should be served under this premise and to determine. The procedure of calculation is shown below.
[0042] (1) Define (the gain change in subsequent games) of a formula (1) as follows.
[Mathematical formula 2] (Gain change in subsequent games) = (gain vector value of the game at the next time) - (gain vector value of a game at present)
(2) Presume the coefficient of expectation w. This value is calculated by trial and error in the process performed by repeating a game.
(3) Calculate the gain vector of the game at the next time using a change rule.
(4) Action is determined using action selection method] which incorporated evaluation of change of the game after being based on [action.
(5) If at least one side will stop a game, it will end there.
It returns to (6) and (3).
[0043] It is assumed that the coefficient of expectation w= 10 was given and each knows Computers A and B about this procedure (2) here. In the procedure (4) which showed the example of the procedure (3) by 53 of drawing 5 , when Computer A takes action a2, no computers B act. Here, it calculates noting that dummy action which nothing carries out is performed. About an initial state, the example which calculated the use is shown in drawing 7 . In an initial state, it turns out that the proposal method chooses action a1 and b1.
[0044] Next, the proposal method is evaluated. The total N of a move cannot be known before a game start. Therefore, the proposal method is compared using the method of making action selection only from the gain vector of a game tree at present. Let a valuation basis be the average of acquisition gain.
[0045]
[Mathematical formula 3] (Average of acquisition gain) =(total of old acquisition gain)/(the number of repetitions of a game)
The number of moves changes with the actions taken among 1-2 about 1 time of a game. Therefore, a term called the number of repetitions of a game was used instead of the number of moves here.
[0046] The number of repetitions of a game and the relation of acquisition gain are shown in drawing 8 . Action a1 and b1 were chosen by the method of this invention each time, and action a2 was chosen by the conventional method each time. A figure shows the following things. By the conventional method, Computer A cannot obtain gain at all. Since Computer B takes the concerted action b1 by the method of this invention to it, gain acquisition of A is possible. Moreover, if the load hung down even if it accompanies tasking of A by taking the concerted action b1 becomes small and the number of repetitions increases also about B, gain

will become large rather than the conventional method. If drawing 8 (b) is seen, the reverse point of the average of gain has happened by 11 repetitions. If a game tree is repeated first and it develops from this to several 11 or more, even if it uses a definition called utility = gain not using a formula (1), it turns out that the same result as the method of this invention can be drawn by the method of determining action strictly backward from the last move. However, 11 or more game trees with a number of repetitions become quite large, and chain calculation time is needed until it takes the first action. In fields, such as control of a physical system, this serves as big demerit and the predominancy of the method of this invention is shown.

[0047]

[Effect of the Invention] [according to this invention] by introducing the description under the change rule of gain to the decision problem of the task request and task undertaking action by two or more decision-making subjects as explained above [evaluating the change condition of the gain in the game which the amount of description of the correspondence relation between an action group and a gain vector is reduced, and is performed henceforth, and including it in the valuation basis of action selection] Even if it loses as in the short term as decision making in a short time, selection of the action which it gains in the long run is attained, and it is effective in making the gain which can be gained in the long run increase.

[Brief Description of the Drawings]

[Drawing 1] It is the block diagram showing the composition of the equipment which enforces the task request in two or more decision-making subject environment and the task undertaking action determination method concerning one embodiment of this invention.

[Drawing 2] It is the figure showing the style of the change rule of a gain vector.

[Drawing 3] It is the figure showing the example computation of the gain vector using the change rule of the gain vector.

[Drawing 4] It is the figure showing the component of the gain vector in an example.

[Drawing 5] It is the figure showing the game tree expression in the initial state of an example, and game tree expression in the following state.

[Drawing 6] It is the figure showing the example computation of the gain vector in an example.

[Drawing 7] It is the figure showing the example computation of the utility vector in an example.

[Drawing 8] It is the figure showing the comparison of the conventional method about change of acquisition gain, and the method of this invention to the number of repetitions of a game.

[Drawing 9] It is the figure showing the example of expression of the game by a game tree.

[Drawing 10] It is the figure showing the composition of the equipment by the conventional

technology.

[Explanations of letters or numerals]

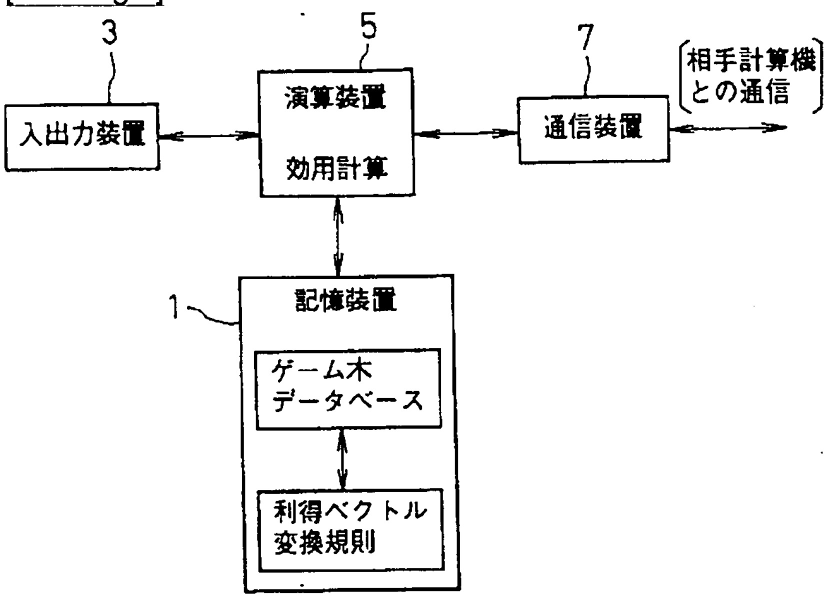
1 Storage

3 I/O Device

5 Arithmetic Unit

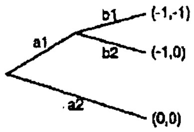
7 Communication Device

[Drawing 1]



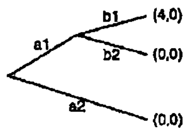
[Drawing 4]

(a)



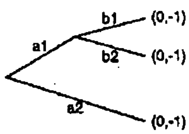
（質問と回答に要するコスト）

(b)



（データ獲得による利得）

(c)



（タスク処理に付随して生じるコスト（初期状態））

a1: 計算機Aがデータ参照依頼を行う
a2: 計算機Aがデータ参照依頼を行わない
b1: 計算機Bがデータ参照依頼に応じる
b2: 計算機Bがデータ参照依頼に応じない

[Drawing 2]

(a)

	$b1$	$b2$
$a1$	$x+1$ $x+2$ $x-1$ $x+1$	$x+0$ $x+0$ $x+0$ $x+0$
$a2$	$x+0$ $x+0$ $x+0$ $x+0$	$x+0$ $x+1$ $x-1$ $x+1$

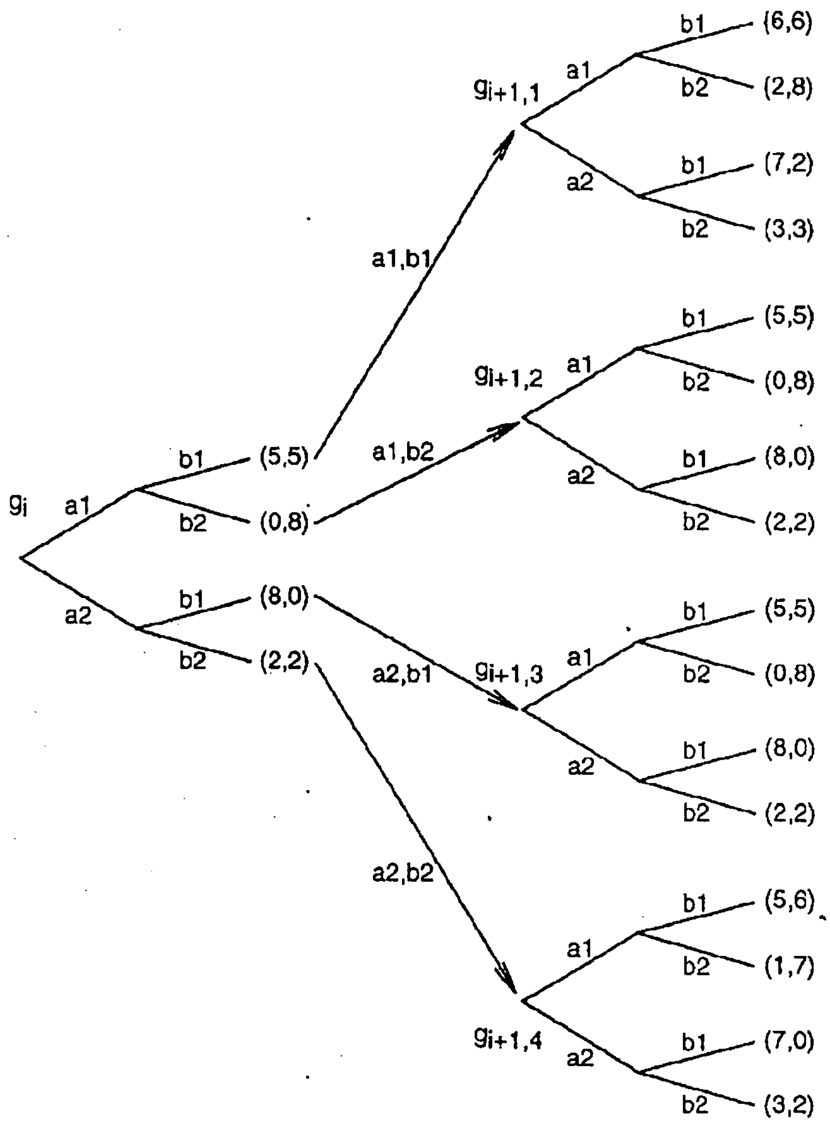
（Aに対する変化規則）

(b)

	$b1$	$b2$
$a1$	$x+1$ $x+0$ $x+2$ $x+1$	$x+0$ $x+0$ $x+0$ $x+0$
$a2$	$x+0$ $x+0$ $x+0$ $x+0$	$x+1$ $x-1$ $x+0$ $x+0$

（Bに対する変化規則）

[Drawing 3]



[Drawing 6]

A:

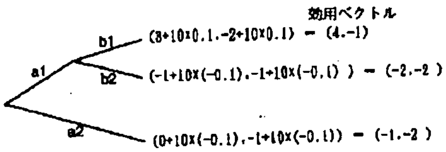
	$b1$	$b2$
$a1$	$f(x+.1)$ $f(x+.1)$ $f(x+.1)$	$f(x-.1)$ $f(x-.1)$ $f(x-.1)$
$a2$	$f(x-.1)$ $f(x-.1)$ $f(x-.1)$	

B:

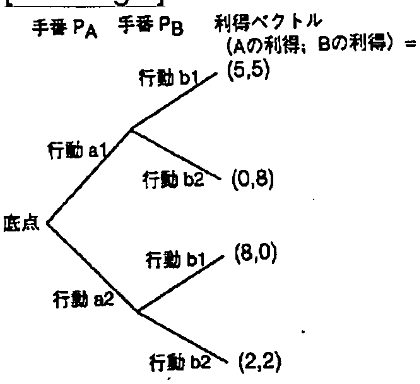
	$b1$	$b2$
$a1$	$f(x+.1)$ $f(x+.1)$ $f(x+.1)$	$f(x-.1)$ $f(x-.1)$ $f(x-.1)$
$a2$	$f(x-.1)$ $f(x-.1)$ $f(x-.1)$	

$$f(x) = \begin{cases} 0 & (x \geq 0\text{のとき}) \\ x & (-2 \geq x < 0\text{のとき}) \\ -2 & (x < -2\text{のとき}) \end{cases}$$

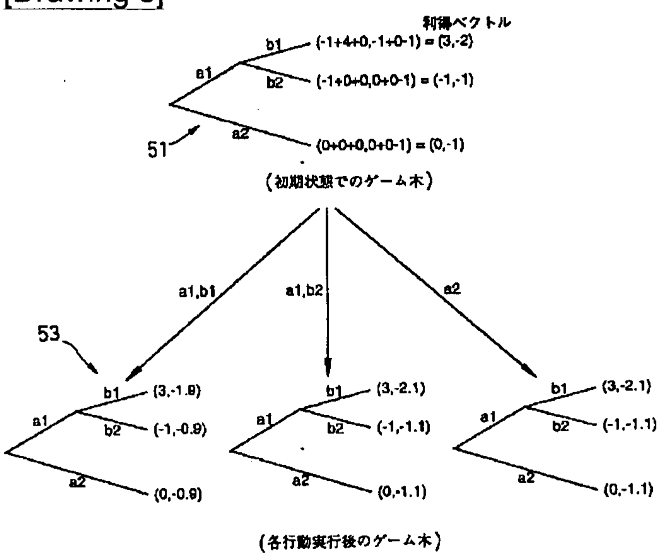
[Drawing 7]



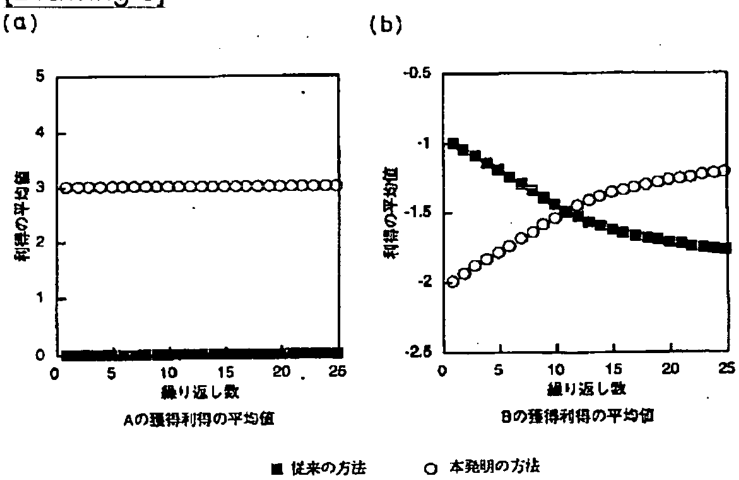
[Drawing 9]



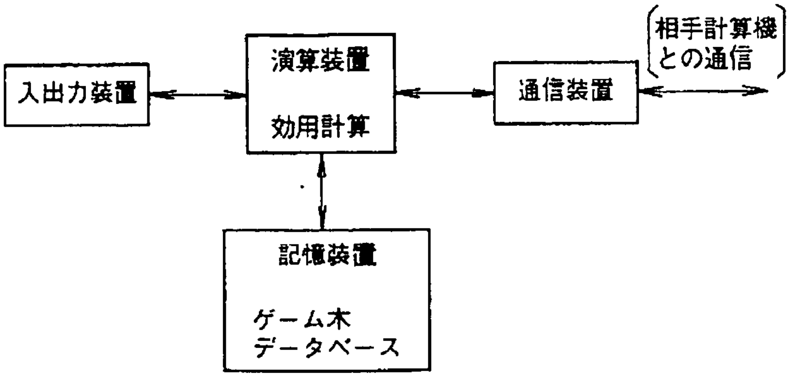
[Drawing 5]



[Drawing 8]



[Drawing 10]



[Translation done.]

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（Ａ）　(11)特許出願公開番号

特開平9－297690

(43)公開日　平成９年(1997)11月18日

(51)Int.Cl.⁶	識別記号	庁内整理番号	ＦＩ		技術表示箇所
Ｇ０６Ｆ　9/46	３６０		Ｇ０６Ｆ　9/46	３６０Ｆ	
15/16	４３０		15/16	４３０Ｂ	

審査請求　未請求　請求項の数２　ＯＬ（全 ８ 頁）

(21)出願番号　特願平8－109724

(22)出願日　平成８年(1996)４月30日

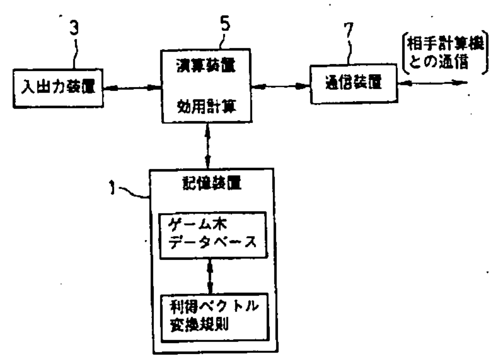
(71)出願人　000004226
日本電信電話株式会社
東京都新宿区西新宿三丁目19番２号

(72)発明者　松原　繁夫
東京都新宿区西新宿三丁目19番２号　日本電信電話株式会社内

(74)代理人　弁理士　三好　秀和　（外１名）

(54)【発明の名称】　複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法

(57)【要約】
【課題】　利得の変化規則による記述を導入することにより短時間で意思決定を行うとともに、長期的には得する行動の選択が可能となり、長期的に獲得できる利得を増加し得る複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法を提供する。
【解決手段】　行動による利得の変化の仕方を変化規則を用いて表現し、必要に応じてその変化規則を用いて将来の状態に対する行動の組合せと利得の関係を計算して導くことにより問題表現を簡便なものとし、またどの行動を取ればよいか計算するときに、現時点での行動が引き起こす将来の利得の変化具合を評価し、その評価を動作選択の判断基準に組み入れることにより、各意思決定主体の長期的利得を増加させる。



【特許請求の範囲】

【請求項１】　複数の意思決定主体が存在し、意思決定主体がそれぞれ複数の行動選択肢を持ち、それぞれの意思決定主体の取った行動の組合せにより各自の利得が決定し、その利得が大きくなるように行動を決定する複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法において、行動の組合せに対する利得を記述するときに、意思決定主体の状態すべてに対して行動の組合せと利得の関係を網羅的に記述することをせずに、行動による利得の変化の仕方を変化規則を用いて表現し、必要に応じてその変化規則を用いて将来の状態に対する行動の組合せと利得の関係を計算して導くことにより問題表現を簡便なものとし、またどの行動を取ればよいか計算するときに、将来の行動から後向き推論して行動を決定せず、現時点の短期的な損得だけから行動を決定することもせず、現時点での行動が引き起こす将来の利得の変化具合を評価し、その評価を動作選択の判断基準に組み入れることにより、各意思決定主体の長期的利得を増加させることを特徴とする複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法。

【請求項２】　複数台の計算機がネットワークを介して接続されたシステムにおける複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法において、計算機の記憶装置に格納された意思決定主体の現時点での行動に対する利得と利得の変化規則とから現時点より後での行動に対する利得を計算する工程と、現時点での意思決定主体の行動による利得とその行動によって引き起こされる将来の利得変化とから計算される効用が最大となるように意思決定主体の行動を決定する工程とを有することを特徴とする複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、相互接続された計算機のネットワークシステム上での、タスク依頼とタスク引き受けの決定問題に対して、処理すべきタスクを持つ計算機が他計算機にデータ転送要求やデータベースの参照要求を行うか自己でタスクを処理するかの行動を決定し、または他計算機からタスク依頼があるときに、それを引き受けるか引き受けないかの行動を決定する複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法に関する。

【０００２】

【従来の技術】複数の計算機がネットワークに接続され、それぞれの計算機がタスクを持ち、タスク処理のために計算機間で情報がやりとりされる環境を考える。これら複数の計算機には全体で１つのタスクが与えられるのではなく、またお互いに協力的行動を取るように指示する主計算機もないとする。ここで、例えば計算機Ａのタスク処理に、計算機Ｂの管理するデータベース内のデータが有効であるとする。ただし、このデータなしでもタスク処理は可能とする。このとき、計算機Ａは計算機Ｂにデータ参照を依頼してタスク処理を行うか、依頼せず自己でタスク処理を行うかの選択がある。また、計算機Ｂは依頼があった場合に、それに応じる応じないの選択がある。

【０００３】一般には、自己の利益となる行動と相手の利益となる行動が常に一致するとは限らない。つまり、自己にとって有利な行動を相手計算機に強制することは不可能である。そのため、相手計算機が合理的に振る舞うという仮定のもとで、自己の利益を最大とする行動を選択しなければならない。このように相手がいる場合の行動選択問題を扱う方法としてゲーム理論がある。

【０００４】ゲーム理論は（１）複数の意思決定主体が存在し、（２）意思決定主体がそれぞれ行動の選択肢を複数持ち、（３）各意思決定主体が行動を選択したとき、その行動の組合せにより、それぞれの利得が決定し、（４）他の意思決定主体は全員、効用（利得）を最大化するように行動するという前提のもとで、どの行動を選択すればよいか決定する方法を与える。また、効用の設定の変更が系全体の性質をどう変化させるかを解析する方法を与えるものである（［鈴木光男、“新ゲーム理論、”勁草書房、1994］、［西田俊夫、“ゲームの理論、”日科技連、1973］参照）。

【０００５】行動選択とそれによる利得の獲得の過程をゲームと呼び、意味決定主体を参加者と呼ぶ。また、ゲームを構成する個々の段階、すなわち、個々の行動選択機会を手番と呼ぶ。ここでは、参加者数が２で、非協力、完全情報のゲームを取り扱う。非協力とは強制力のある事前の取り決めが存在しないことをいう。結果として協調行動を取ったように見えるか見えないかは全く関係がない。また、完全情報とは、すべての参加者がそれ以前の過程で選択された行動をすべて知っていることをいう。

【０００６】ゲームを表す方法の１つとして、ゲーム木による表現がある。ゲーム木によるゲームの表現例を図９に示す。P_A，P_B はそれぞれゲームの参加者Ａ，Ｂの手番を表す。枝は選択可能な行動を表す。この図は、Ａが最初にａ１かａ２の行動を選択し、それを見てからＢがｂ１かｂ２の行動を選択する過程を表している。参加者Ａ，Ｂの行動の組合せによりそれぞれの利得が決まり、それが端点に示されている。これを利得ベクトルと呼ぶ。

【０００７】さて、ゲームは双方が１回ずつ行動を選択して終るだけでなく、図９に示したゲームを部分として繰り返すような、何段階も手が続くゲームがある。これを多段階ゲームという。本明細書では、現時点から双方が１手ずつ行動を選択して、最初に利得が確定するまでを１つの単位として現時点のゲームと呼ぶ。

【０００８】さて、手番の総数がＮとなるゲーム木が与

えられたとき、各手番で行動選択する従来の方法を以下に示す。ここでは、効用＝利得と定義し、Nを偶数とする。
【0009】（ゲーム木上での行動選択方法）
（1）ｉ←Nとする。
（2）手番ｉの各分岐点について、効用ベクトルを見て、この手番の参加者の効用を最大にする行動を各分岐点での選択行動する。ここでは、効用ベクトルは利得ベクトルと同一である。
（3）手番ｉ－１の各分岐点について、（2）での選択行動に対する効用ベクトルの中から、この手番の参加者の効用を最大にする行動を選んで、各分岐点での選択行動とする。
（4）ｉ＝２であれば、ｉ←（ｉ－１）として（7）へ行く。
（5）ｉ＝２でなければ、分岐点ｉ－２での各行動に対する効用ベクトルに、それに後続する手番ｉ－１での選択行動に対する効用ベクトル値を加える。
（6）ｉ←（ｉ－２）として、（2）へ行く。
（7）手番ｉで選択された分岐点での選択行動をこのゲームでの選択行動とする。手番１では分岐点は１つであり、それを選択された分岐点とする。
（8）ｉ＝Nであれば、終了する。
（9）（7）での選択行動に後続する分岐点をその手番ｉ←（ｉ＋１）での選択分岐点とする。
（10）（7）へ行く。
【0010】完全情報ゲームでは、各参加者はどの分岐点にいるか知ることができ、また各分岐点で参加者は常にそれ以降で自己の効用を最大化する行動を選択するという事実から、上の方法の非協力ゲームにおける正当性が導ける。
【0011】図9の例を参照すると、手番２（P_B）では、どちらの分岐点でも、Bにとって、行動ｂ２の効用がｂ１よりも大きいので、行動ｂ２が選択される。次に、手番１（P_A）では、ａ１（効用ベクトル（0，8））とａ２（効用ベクトル（2，2））の比較になる。Aにとっては、ａ２の方が効用が大きいのでこちらが選択される。結局このゲームでは、行動系列ａ２→ｂ２が現れる。
【0012】
【発明が解決しようとする課題】さて、利得ベクトルの入力処理を人間が行い、行動選択の計算を計算機が行うとする。従来技術を用いて実現する場合の構成図を図10に示す。ゲームをゲーム木として表現しようとするとき、以下の問題が生じる。
【0013】（ａ）一般には、現時点で取る行動がそれ以降行うゲームの利得ベクトルを変化させる。そのため、その変化を取り扱えることが必要である。大きさの小さいゲームを１つだけ用意して、それを繰り返し行うという扱いでは不十分である。
【0014】（ｂ）手番の数Nが大きくなると、ゲーム木が膨大となるため、予め展開した形に書き下しておくことが困難となる。
【0015】（ｃ）参加者の状態とそこで行われるゲームの利得ベクトルの組を記述しておけば、必要に応じてゲーム木を展開することができる。しかし、状態数が増えれば、状態と利得ベクトルの組の記述量が膨大となり、管理が困難となる。
【0016】また、計算を行う上で、以下の問題が生じる。
【0017】（ｄ）前述の行動選択に関する従来方法は最後の手番Nから、まず後向きに計算が進行する。しかし、参加者をゲームに参加させ続ける強制力はどの参加者にもないため、予めどれだけの手番が実行されるか知ることはできない。そのため、Nが与えられない場合、どこから計算を始めればよいか決定できない。
【0018】（ｅ）手番の総数Nがわかっている場合でも、Nが大きくなると、ゲーム木が膨大になり、計算量が増加し、行動決定までに時間を要するようになる。
【0019】（ｆ）最後の手番Nからの計算を諦めて、現時点のゲームに関して行動決定する場合、将来の損得を計算に含めることができない。そのため、短期的には損をしても長期的に利得を得る行動の選択が不可能になる。
【0020】本発明は、上記に鑑みてなされたもので、その目的とするところは、利得の変化規則による記述を導入することにより短時間で意思決定を行うとともに、長期的には得する行動の選択が可能となり、長期的に獲得できる利得を増加し得る複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法を提供することにある。
【0021】
【課題を解決するための手段】上記目的を達成するため、請求項１記載の本発明は、複数の意思決定主体が存在し、意思決定主体がそれぞれ複数の行動選択肢を持ち、それぞれの意思決定主体の取った行動の組合せにより各自の利得が決定し、その利得が大きくなるように行動を決定する複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法において、行動の組合せに対する利得を記述するときに、意思決定主体の状態すべてに対して行動の組合せと利得の関係を網羅的に記述することをせずに、行動による利得の変化の仕方を変化規則を用いて表現し、必要に応じてその変化規則を用いて将来の状態に対する行動の組合せと利得の関係を計算して導くことにより問題表現を簡便なものとし、またどの行動を取ればよいか計算するときに、将来行動から後向き推論して行動を決定せず、現時点の短期的な損得だけから行動を決定することもせず、現時点での行動が引き起こす将来の利得の変化具合を評価し、その評価を動作選択の判断基準に組み入れることにより、各意思決定主

体の長期的利得を増加させることを要旨とする。

【0022】請求項1記載の本発明にあっては、行動による利得の変化の仕方を変化規則を用いて表現し、必要に応じてその変化規則を用いて将来の状態に対する行動の組合せと利得の関係を計算して導くことにより問題表現を簡便なものとし、またどの行動を取ればよいか計算するときに、現時点での行動が引き起こす将来の利得の変化具合を評価し、その評価を動作選択の判断基準に組み入れることにより、各意思決定主体の長期的利得を増加させる。

【0023】また、請求項2記載の本発明は、複数台の計算機がネットワークを介して接続されたシステムにおける複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法において、計算機の記憶装置に格納された意思決定主体の現時点での行動に対する利得と利得の変化規則とから現時点より後での行動に対する利得を計算する工程と、現時点での意思決定主体の行動による利得とその行動によって引き起こされる将来の利得変化とから計算される効用が最大となるように意思決定主体の行動を決定する工程とを有することを要旨とする。

【0024】請求項2記載の本発明にあっては、意思決定主体の現時点での行動に対する利得と利得の変化規則とから現時点より後での行動に対する利得を計算し、現時点での意思決定主体の行動による利得とその行動によって引き起こされる将来の利得変化とから計算される効用が最大となるように意思決定主体の行動を決定する。

【0025】

【発明の実施の形態】以下、図面を用いて本発明の実施の形態について説明する。

【0026】図1は、本発明の一実施形態に係る複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法を実施する装置の構成を示すブロック図である。同図に示す装置は、記憶装置1、入出力装置3、効用計算を行う演算装置5、相手計算機と通信を行う通信装置7を有し、記憶装置1には図10に示した従来の装置に＊対して利得ベクトルの変化規則をゲーム木のデータベースに加えて新たに追加した点が異なるものである。このように利得ベクトルの変化規則を別に記述する方法を新たに追加することにより、ゲームのゲーム木による表現を容易にすることができる。これは、現時点のゲームを行い、ある行動を実行すると、それ以降のゲームで利得ベクトルがどう変化するかを記述するものである。

【0027】図2は、図9に対する変化規則の記述例を示しているものである。図2(a)はAに対する変化規則であり、図2(b)はBに対する変化規則である。また、同図のテーブルにおいて、xはそれ以前のゲームにおける利得の値を示す。a1，b1という行動が取られた場合、テーブルの行a1と列b1の交差部を見る。Aについては、式x+1，x+2，x−1，x+1とあり、それぞれ、次の時点での(a1，b1)，(a1，b2)，(a2，b1)，(a2，b2)の利得の計算式に対応する。実際に計算すると、以下のように求まる。

【0028】(a1，b1)：　x+1=5+1=6

(a1，b2)：　x+2=0+2=2

(a2，b1)：　x−1=8−1=7

(a2，b2)：　x+1=2+1=3

【0029】同様に他の行動組やBに対して計算すると、図3が得られる。必要に応じて、以降のゲームの利得ベクトルも計算できる。

【0030】次に、現時点での行動と以降のゲームの関係の評価を加えた行動選択方法を提案する。これは、現時点での行動が、それ以降で、自己にとってより有利なゲーム、すなわちより大きな利得を得る可能性のあるゲームに導くか導かないかを評価し、それと現時点でのゲーム上の損得の評価とを合わせて、行動選択を行うものである。効用を以下の式で定義する。

【0031】

【数1】

$$\text{効用}=(\text{現時点でのゲームの利得})+(\text{期待係数}w)\times(\text{以降のゲームでの利得変化}) \qquad (1)$$

ゲームの参加者はこの効用を最大にする行動を選択する。期待係数は、参加者毎に異なり、それぞれの持つタスクに応じて調整される。

【0032】提案方法の計算手順を以下に示す。ここで、現時点のゲームとは双方が1手ずつ行動を選択して、利得が確定するまでのゲームを指すことを注意しておく。

【0033】(現時点の行動による以降のゲーム変化の評価を組み入れた行動選択方法)

(1) 現時点の手番をiとする。

(2) 現時点のゲームの利得から上式(1)の効用を計算する。

(3) 手番i+1の各分岐点について、この手番の参加※者の効用を最大にする行動を各分岐点での選択行動とする。

(4) 手番iの各分岐点について、(3)での選択行動に対する効用ベクトルの中から、この手番の参加者の効用を最大にする行動を選んで、このゲームでの選択行動とする。

(5) 分岐点iでの選択行動に後続する分岐点での選択行動を手番i+1での選択行動とする。

(6) 終了。

【0034】上記の変化規則による表現の導入により、状態と利得ベクトルの組だけの記述よりも、記述量の削減が期待できる。必要に応じてゲーム木を展開することにすれば、必要記憶容量を削減できる。

【0035】また、上記の行動による以降のゲーム変化の評価を組み入れた行動選択方法により以下のことが可能となる。

【0036】（ａ）提案手法は従来技術のように厳密な計算を行わず、将来の利得の増分を推定するため、常に正しい判断ができるとは限らない。しかし、従来技術のようにゲーム木全体を網羅的に計算する必要がなくなり、計算時間が削減される。これは、物理システムの制御など時間的制約がある場合特に有効である。

【0037】（ｂ）手番の総数Ｎが予めわからない場合でも、現時点の損得だけでの判断ではなくなり、現時点で損をしても、将来大きな利益を得る可能性がある行動があれば、それを選択する可能性が生じる。

【0038】（ｃ）期待係数の操作により、参加者の現時点での負荷に応じた適切な行動選択が可能となる。例えば、現時点で余裕があれば、期待度を大きくして、相手からの負荷の高い依頼に協調的に応じることができる。相手も同じ行動選択方法を用いれば、自己の負荷が高く、相手にタスクを分配したいときに、相手が承諾する可能性が高まる。また、余裕がなければ、期待係数を小さくして相手からの依頼を拒絶し、自己のタスク処理に専念することができる。

【0039】次に、具体的な例として、計算機Ａと計算機Ｂとの間で情報流通問題を考える。ここでの問題設定を以下に示す。

（１）計算機Ａと計算機Ｂは同一ネットワークに接続している。

（２）計算機Ａはタスク処理のため、相手計算機ＢにＢが管理するデータベースのデータ参照要求を出す。

（３）計算機Ａは相手に参照依頼を出すとコスト（負の利得）が発生する。

（４）計算機Ｂは相手計算機のデータベース参照要求に応じるとコスト（負の利得）が発生する。相手計算機の要求を拒絶すれば、コストはかからない。

（５）計算機Ａは相手からデータを獲得できるとそれがタスク処理に役立つため正の利得を得ることができる。

（６）計算機Ａはタスク処理時にネットワークに不要なパケットを流すなどして、相手計算機に負荷をもたらす。負荷の程度は計算機Ａの知識レベルによる。ネットワーク上での情報流通に関する知識レベルが高いときには、周囲に与える負荷を小さくできる。

（７）計算機Ｂが参照要求に応じると、それに付随する情報流通に関する知識が計算機Ａに蓄積され、Ａの知識レベルが上がる。

（８）計算機間で情報のやりとりがない場合、計算機Ａは情報流通に関する知識を得ることができず、知識レベルが下がる。

【0040】この問題の行動と利得の関係を図4に示す。質問と回答に関するコストを図4（ａ）に示す。データ獲得による利得を図4（ｂ）に示す。タスク処理に付随してＢに生じるコストの初期値を図4（ｃ）に示す。このゲームを１回行うことによる利得は、（ａ），（ｂ），（ｃ）の和となる。初期状態でのゲーム木を図5の51で示す。図4（ｃ）のコストは、計算機Ａの知識レベルにより変化する。そのコストの変化規則を図6に示す。この規則を用いて求めた初期状態の次の状態でのゲーム木を図5の53で示す。

【0041】この前提のもとで、計算機Ａ，Ｂはどのように振る舞えばよいか決定することが、ここでの問題である。計算の手順を以下に示す。

【0042】（１）式（１）の（以降のゲームでの利得変化）を以下のように定義する。

【数2】（以降のゲームでの利得変化）＝（次の時点のゲームの利得ベクトル値）－（現時点のゲームの利得ベクトル値）

（２）期待係数ｗを推定する。この値は、ゲームを繰り返し行う過程で試行錯誤的に求められる。

（３）次の時点のゲームの利得ベクトルを変化規則を用いて計算する。

（４）［行動による以降のゲームの変化の評価を組み入れた行動選択方法］を用いて行動を決定する。

（５）少なくとも一方がゲームを止めることにすれば、そこで終了する。

（６）（３）へ戻る。

【0043】この手順（２）について、ここでは、計算機Ａ，Ｂとも期待係数ｗ＝１０が与えられ、お互いが知っているとする。手順（３）の例は図5の５３で示した手順（４）において、計算機Ａが行動ａ２を取った場合、計算機Ｂは何も行動しない。ここでは、何もしないダミー行動を行うとして計算を行う。初期状態について、効用を計算した例を、図7に示す。初期状態では、提案方法は行動ａ１，ｂ１を選択することがわかる。

【0044】次に、提案方法の評価を行う。ゲーム開始前に手番の総数Ｎを知ることはできない。そのため、現時点のゲーム木の利得ベクトルだけから行動選択する方法を用いて、提案方法を比較する。評価基準は獲得利得の平均値とする。

【0045】

【数3】（獲得利得の平均値）＝（これまでの獲得利得の総和）／（ゲームの繰り返し数）

１回のゲームにつき、取られる行動により手番の数が１～２の間で変化する。そのため、ここでは手番の数の代わりにゲームの繰り返し数という用語を用いた。

【0046】図8にゲームの繰り返し数と獲得利得の関係を示す。本発明の方法では、行動ａ１，ｂ１が毎回選択され、従来の方法では、行動ａ２が毎回選択された。図から以下のことがわかる。従来方法では計算機Ａは全く利得を得ることができない。それに対し、本発明の方法では、計算機Ｂが協調行動ｂ１を取るため、Ａの利得獲得が可能となっている。また、Ｂについても、協調行

動ｂ１を取ることにより、Ａのタスク処理に付随してもたらされる負荷が小さくなり、繰り返し数が増えると従来方法よりも利得が大きくなる。図８（ｂ）を見ると、繰り返し数１１で利得の平均値の逆点が起こっている。このことから、ゲーム木を最初に繰り返し数１１以上まで展開しておけば、式（１）を用いず、効用＝利得という定義を用いても、最後の手番から後向きに厳密に行動を決定する方法により本発明の方法と同様の結果が導けることがわかる。しかし、繰り返し数１１以上のゲーム木はかなり大きくなり、最初の行動を起こすまでにより長い計算時間が必要となる。物理システムの制御などの分野ではこれは大きな短所となり、本発明の方法の優位性が示される。

【００４７】

【発明の効果】以上説明したように、本発明によれば、複数の意思決定主体によるタスク依頼とタスク引き受け行動の決定問題に対し、利得の変化規則による記述を導入することにより、行動組と利得ベクトルの対応関係の記述量が削減され、また以降に行うゲームでの利得の変化具合を評価し、それを行動選択の評価基準に含めることで、短時間での意思決定と短期的には損をしても長期的には得をする行動の選択が可能となり、長期的に獲得できる利得を増加させる効果がある。

【図面の簡単な説明】

【図１】本発明の一実施形態に係る複数意思決定主体環境でのタスク依頼とタスク引き受け行動決定方法を実施する装置の構成を示すブロック図である。

【図２】利得ベクトルの変化規則の表現法を示す図である。

【図３】利得ベクトルの変化規則を用いた利得ベクトルの計算例を示す図である。

【図４】例題における利得ベクトルの構成要素を示す図である。

【図５】例題の初期状態におけるゲーム木表現と次の状態でのゲーム木表現を示す図である。

【図６】例題における利得ベクトルの計算例を示す図である。

【図７】例題における効用ベクトルの計算例を示す図である。

【図８】ゲームの繰り返し数に対する獲得利得の変化についての従来方法と本発明の方法の比較を示す図である。

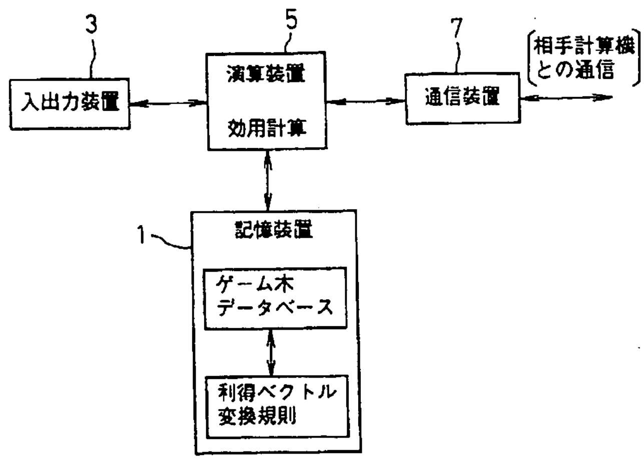
【図９】ゲーム木によるゲームの表現例を示す図である。

【図１０】従来技術による装置の構成を示す図である。

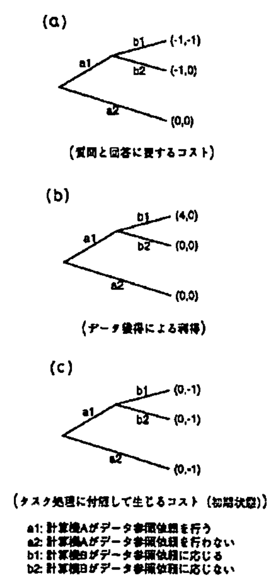
【符号の説明】

１　記憶装置
３　入出力装置
５　演算装置
７　通信装置

【図１】



【図４】



【図2】

(a)

	b1	b2
a1	$x+1$　$x+2$ $x-1$　$x+1$	$x+0$　$x+0$ $x+0$　$x+0$
a2	$x+0$　$x+0$ $x+0$　$x+0$	$x+0$　$x+1$ $x-1$　$x+1$

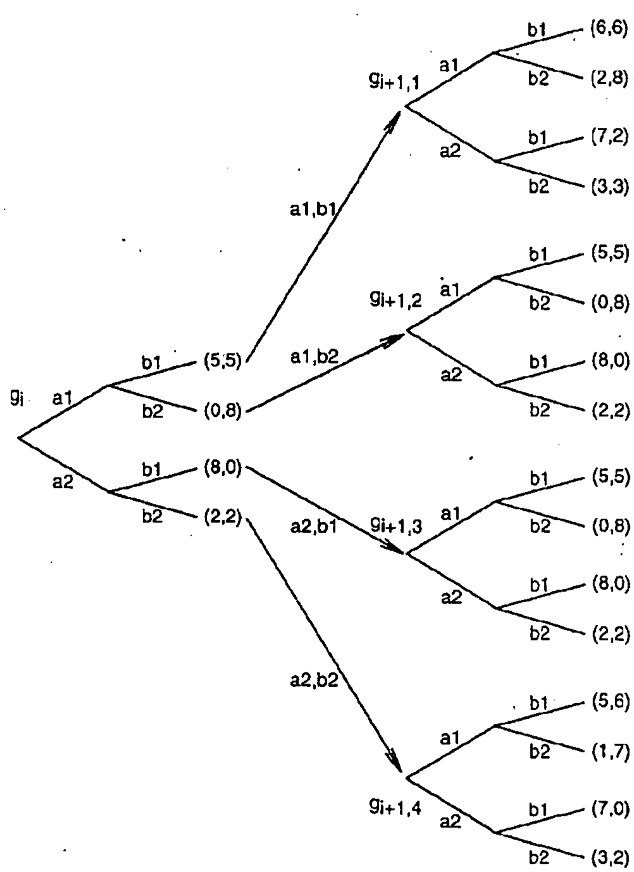
（Aに対する変化規則）

(b)

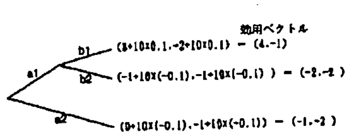
	b1	b2
a1	$x+1$　$x+0$ $x+2$　$x+1$	$x+0$　$x+0$ $x+0$　$x+0$
a2	$x+0$　$x+0$ $x+0$　$x+0$	$x+1$　$x-1$ $x+0$　$x+0$

（Bに対する変化規則）

【図3】



【図7】



【図6】

A :

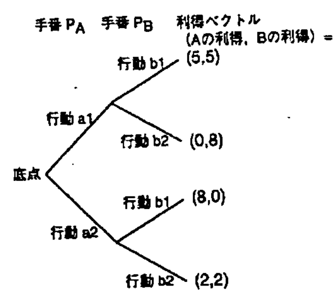
	b1	b2
a1	$f(x+.1)$　$f(x+.1)$ $f(x+.1)$	$f(x-.1)$　$f(x-.1)$ $f(x-.1)$
a2	$f(x-.1)$　$f(x-.1)$ $f(x-.1)$	

B :

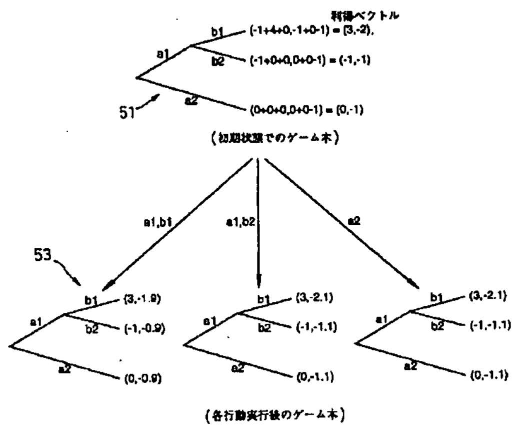
	b1	b2
a1	$f(x+.1)$　$f(x+.1)$ $f(x+.1)$	$f(x-.1)$　$f(x-.1)$ $f(x-.1)$
a2	$f(x-.1)$　$f(x-.1)$ $f(x-.1)$	

$$f(x)=\begin{cases}0 & (x\geq 0\text{のとき})\\ x & (-2\geq x<0\text{のとき})\\ -2 & (x<-2\text{のとき})\end{cases}$$

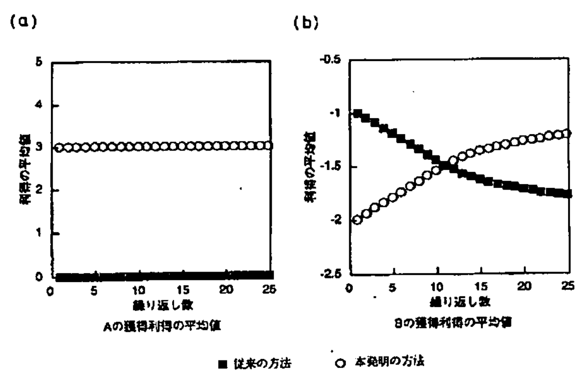
【図9】



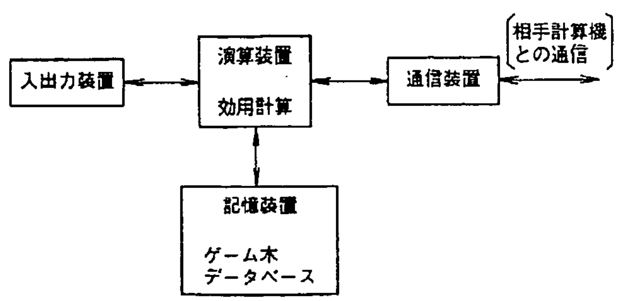
【図5】


利得ベクトル
b1
(-1+4+0,-1+0-1) = (3,-2),
a1
b2
(-1+0+0,0+0-1) = (-1,-1)
51
a2
(0+0+0,0+0-1) = (0,-1)
（初期状態でのゲーム木）
a1,b1
a1,b2
a2
53
b1
(3,-1.9)
a1
b2
(-1,-0.9)
a2
(0,-0.9)
b1
(3,-2.1)
a1
b2
(-1,-1.1)
a2
(0,-1.1)
b1
(3,-2.1)
a1
b2
(-1,-1.1)
a2
(0,-1.1)
（各行動実行後のゲーム木）

【図8】


(a)
(b)
利得の平均値
繰り返し数
Aの獲得利得の平均値
Bの獲得利得の平均値
■ 従来の方法
○ 本発明の方法

【図10】


入出力装置
演算装置
効用計算
通信装置
相手計算機との通信
記憶装置
ゲーム木
データベース